

Closely related languages, different ways of realizing focus

Szu-wei Chen¹, Bei Wang², Yi Xu³

¹ Graduate Institute of Linguistic, National Chung Cheng University, Taiwan

² Institute of Chinese Minority Languages, Minzu University of China, China

³ Department of Speech, Hearing and Phonetic Sciences, University College London, UK

chen.szuwei@gmail.com, bjwangbei@googlemail.com, yi.xu@ucl.ac.uk

Abstract

We investigated how focus was prosodically realized in Taiwanese, Taiwan Mandarin and Beijing Mandarin by monolingual and bilingual speakers. Acoustic analyses showed that all speakers raised pitch and intensity of focused words, but only Beijing Mandarin speakers lowered pitch and intensity of post-focus words. Cross-group differences in duration were mixed. When listening to stimuli from their own language groups, subjects from Beijing had over 80% focus recognition rate, while those from Taiwan had less than 70% recognition rate. This difference is mainly due to presence/absence of post-focus compression. These findings have implications for prosodic typology, language contact and bilingualism.

Index Terms: focus, language contact, bilingualism

1. Introduction

An important issue about prosodic focus is whether and how it can be realized in a tone language. Lexical tones, which use F_0 as their major acoustic carrier, are in apparent conflict with focus, which is known to also use F_0 variations as its major acoustic correlate [2, 10, 11]. Thus it is possible that the presence of lexical tones in a language would prevent F_0 from being used to encode focus [7]. However, research on Mandarin, a tone language, and Japanese, a pitch accent language, has shown that focus can be realized by F_0 variations that are independent of those due to lexical contrasts. In particular, a common way to realize focus is to not only expand the on-focus pitch range but also compress the post-focus pitch range [6, 8, 10, 13]. This way of realizing focus seems to be highly effective, as suggested by perceptual results [8].

The manner of focus realization in Mandarin as found in previous research [8, 13], might suggest that this is a common feature of the Chinese language family. But a recent study has shown that in Taiwanese, which is a branch of Southern Min Chinese (Min Nan Hua) spoken in Taiwan, duration is a more consistent cue than F_0 for signaling focus [9]. A close inspection of [9], however, shows that the focus related duration pattern in Taiwanese is not very different from that in Beijing Mandarin [13]. Thus it is not yet clear whether and how focus is prosodically realized in Taiwanese. In the present study we tried to answer this question by directly comparing the production and perception of focus in Taiwanese, Taiwan Mandarin and Beijing Mandarin.

Beijing Mandarin is the local dialect of the City of Beijing, but its phonetic system is also the basis of Standard Chinese, the official language of mainland China. Taiwan Mandarin, is a variant of Mandarin spoken in Taiwan. Although once homogeneous with Standard Chinese, at least by definition, it now has noticeable differences in vocabulary, grammar [4]

and pronunciation [5] from its mainland counterpart. Mandarin was strongly promoted by the Nationalist government until the 1980s, and it remains dominant in Taiwan. However, with the largest ethnic group Hoklo, Taiwanese was also spoken widely at home and among friends. Today, most people in Taiwan are bilinguals, fluent in both Taiwanese and Taiwan Mandarin. Over the years, Taiwan Mandarin has acquired many Taiwanese features in both syntax [4] and phonology [17]. However, given the functional importance of focus, we expected little change in focus realization in Taiwan Mandarin from that in Beijing Mandarin. So it could serve as a good control to highlight the difference in focus realization in Taiwanese.

2. Method

2.1. Production experiment

2.1.1. Stimuli

The target sentence is made up of three words consisting of syllables with identical underlying tones (tone 1, high-level) in both Mandarin and Taiwanese, as shown in Table 1. Although the lexical items in this sentence are the same, for Taiwanese, the tone of the 1st, 3rd and 4th syllables changes into tone 7 (mid) due to a tone sandhi rule [3].

Table 1. Target sentences in Taiwanese and Mandarin. The numbers in the transcription indicate the underlying tone

	Word 1	Word 2	Word 3
Characters & gloss	媽媽 “mother”	摸 “stroke”	貓咪 “kitty”
Taiwanese transcription	/ma1ma1/	/boŋ1/	/niau1mi1/
Mandarin transcription	/ma1ma1/	/mo1/	/mao1mi1/

Table 2. Precursor questions in Mandarin for eliciting four types of focus.

Focus	Precursor questions	English translation
None	圖中你看見什麼?	What do you see in the picture?
Initial	誰在摸貓咪?	Who is stroking the kitty?
Medial	媽媽對貓咪做什麼?	What is Mom doing to the kitty?
Final	媽媽在摸什麼?	What is Mom stroking?

To elicit focus on different words in the sentence, a picture illustrating “Mom is stroking the kitty” was prepared. And a set of precursor questions, each asking about a specific aspect of the picture, were used to elicit one of four types of focus: none, initial (on word 1), medial (on word 2), and final (on

word 3), as shown in Table 2. The target sentences and their precursor questions were randomized and repeated five times. Thus there were 4 foci x 5 repetitions = 20 sentences for each language.

2.1.2. Subjects

Four groups of 8 speakers, each with 4 males and 4 females, participated as subjects, as shown below.

Group 1: Monolingual Taiwanese, age 28-58

Group 2: Monolingual Taiwan Mandarin, age 26-60

Group 3: Bilingual Taiwanese and Taiwan Mandarin, age 28-32

Group 4: Monolingual Beijing Mandarin, age 18-30

Each monolingual speaker recorded one set of data, in Mandarin or Taiwanese, while bilingual speakers recorded two sets of data, in both Mandarin and Taiwanese. To guarantee minimal dialectal variability in Taiwanese, only speakers born and raised in Kaohsiung and Tainan were recruited as Taiwanese subjects. None of subjects reported having any speech disorders.

2.1.3. Recording procedure

Each recording session took place in a quiet room. For the recording sessions in Taiwan, the speech was directly digitized into a SONY Hi-MD (MZ-RH1) using a unidirectional microphone (Audio-Technica AT 9470) placed about 5 to 10 inches from the subject's lips. For the recording sessions in Beijing the speech was digitized into a computer by a 24Bit/96K Firewire Recording System (PreSonus Firebox) using condenser microphone (Rode NT1-A). During each trial the experimenter read aloud the precursor question, and the subject read aloud the target sentence as an answer to the question.

2.1.4. Analysis and results

The extraction of F_0 contours was done with a procedure that combines automatic vocal pulse marking by Praat [1] and manual rectification using a custom-written Praat script [14]. The script then generated a locally smoothed F_0 contour for each sentence, and computed mean F_0 , mean intensity and duration of each syllable. Figure 1 displays time-normalized mean F_0 contours produced by all speaker groups.

In Figure 1 the mean F_0 contours of Taiwanese, by both monolingual and bilingual speakers, show very little difference across the 4 focus conditions (Figs. 1a, 1c). Larger differences can be seen in the F_0 contours of Taiwan Mandarin speakers, especially those by monolingual speakers (Figs. 1b, 1d). However, in all these cases post-focus F_0 in initial and medial focus sentences does not go below the F_0 of the corresponding words in the no focus condition. In contrast, post-focus F_0 is substantially lowered in the case of Beijing Mandarin (Figure 1e).

Figure 2 displays the differences in mean F_0 , mean intensity and duration between the on-focus words and the their no-focus counterparts. It can be seen that on-focus raising of F_0 (Fig. 2a), intensity (Fig. 2c) and duration (Fig. 2e) is produced by all speaker groups, and often more by speakers from Taiwan than by those from Beijing. In contrast, only Beijing Mandarin speakers produced post-focus lowering of F_0 and intensity (Fig. 2b, 2d). Two-way (speaker group, focus type) mixed ANOVAs showed significant effect of speaker group on on-focus duration change ($F[4,35] = 4.09, p = 0.008$), but not on on-focus change in mean F_0 or mean intensity. Another set of two-way ANOVAs showed significant effect of

speaker group on post-focus change in mean F_0 ($F[4,35] = 12.32, p < .0001$), mean intensity ($F[4,35] = 3.516, p = 0.0163$) and duration ($F[4,35] = 2.81, p = 0.0401$). But as can be seen in Figure 2f, post-focus duration is lengthened rather than shortened by speakers from Taiwan except monolingual Taiwan Mandarin speakers.

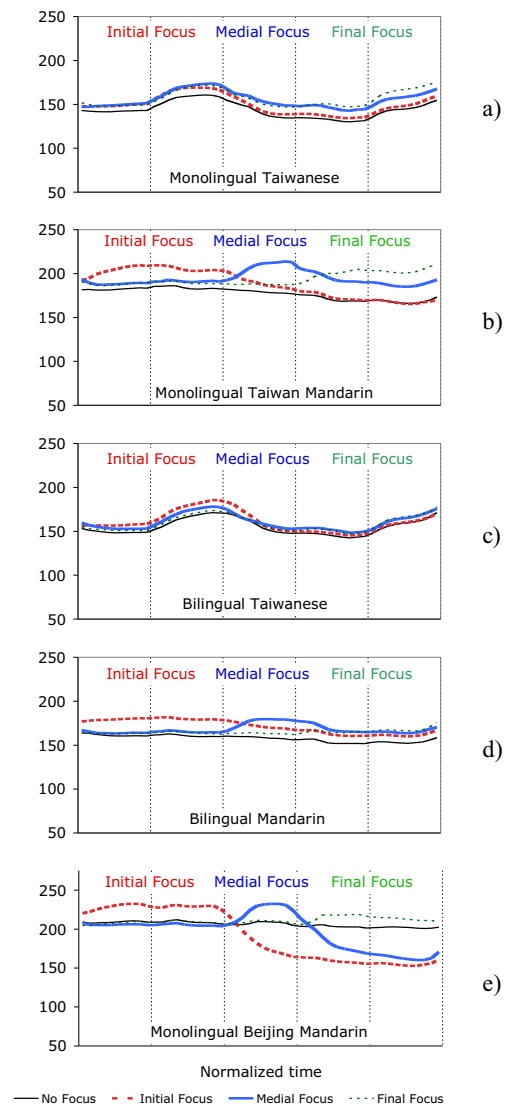


Figure 1: Time-normalized mean F_0 contours produced by 4 speaker groups. Each curve is an average of 40 repetitions by 8 speakers. The vertical lines indicate syllable boundaries. The solid thin curves are from the no-focus condition.

2.2. Perception experiment

2.2.1. Stimuli

All the stimuli came from sentences recorded in Experiment 1. For each language group, three speakers were selected based on their mean standard deviation of all F_0 point across the four focus conditions: those with maximum, minimum or median standard deviations. All 5 tokens recorded for each of these three speakers were used. So, for each language group, there were 4 foci x 5 repetitions x 3 speakers = 60 tokens.

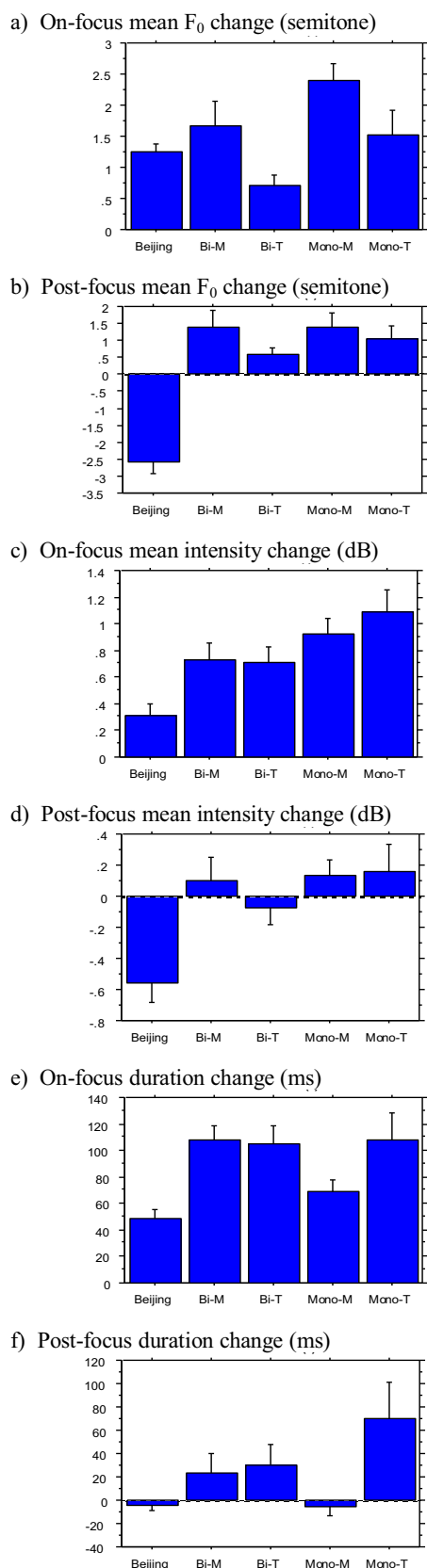


Figure 2: Differences in mean F_0 , intensity and duration between on- and post-focus sentences and the no-focus sentences. Data from Beijing Mandarin (Beijing), bilingual Taiwan Mandarin (Bi-M), bilingual Taiwanese (Bi-T), monolingual Taiwan Mandarin (Mono-M) and monolingual Taiwanese (Mono-T) speakers.

2.2.2. Subjects

Four groups of listeners, as shown below, participated as subjects, each listening to focus samples from their own matched language groups. The bilingual group listened to both Taiwanese and Taiwan Mandarin stimuli produced by bilingual speakers. Listeners had no self-reported speech and hearing disorders.

Group 1: 10 monolingual Taiwanese speakers, 5 females, 5 males, age 46-60.

Group 2: 10 monolingual Taiwan Mandarin speakers, 5 females, 5 males, age 25-40.

Group 3: 10 bilingual speakers, 5 females, 5 males, age 28-52.

Group 4: 11 monolingual Beijing Mandarin speakers, 6 females, 5 males, age 18-23.

2.2.3. Listening procedure

Subjects were asked to listen to the sentence “Mama bong niaumi” (Taiwanese) or “Mama mo maomi” (Mandarin) and judge which of the three words, or none of the words, was emphasized. They were given five practice trials before the real trials *without feedback on the correctness of their answers* so as not to be biased in any way. In each trial the stimulus sentence was played only once. The perception experiment was done using ExperimentMFC in Praat.

2.2.4. Results

Table 3 shows the confusion matrix of focus perception. It can be seen that the overall focus recognition rate is higher for Beijing listeners than for Taiwan listeners. A two-way (speaker group, focus type) mixed ANOVA showed significant effect of speaker group ($F[4,46] = 14.73, p < .0001$), but no effect of focus. There was a significant interaction of speaker group and focus type ($F[12,138] = 14.211, p = 0.0202$). This interaction is due to the fact that the greatest differences between the Beijing and Taiwan listeners are for initial and medial focus as can be seen in Table 3, for which compression of post-focus F_0 and intensity is possible. For final focus and no focus, Beijing listeners did not do much better than the other listeners.

3. Discussion

The acoustic analyses in Experiment 1 and perceptual tests in Experiment 2 demonstrate that there are clear differences in the manner of prosodically realizing focus between Taiwanese and Taiwan Mandarin on the one hand and Beijing Mandarin on the other. Acoustically, the main difference is in terms of the presence and absence of post-focus compression of both F_0 and intensity: In Beijing Mandarin, F_0 and intensity of post-focus words are substantially lowered, while in Taiwanese and Taiwan Mandarin, spoken by both monolingual and bilingual speakers, such post-focus compression is entirely absent. At the same time, all the speakers increased F_0 , intensity and duration of on-focus words. There is virtually no reduction of the duration of the post-focus words, and in fact post-focus duration is *increased* in Taiwanese by both monolingual and bilingual speakers, and in Taiwan Mandarin by bilingual speakers. Thus they increased the duration of all syllables whenever there is a focus anywhere in the sentence, which does not seem to be an effective way of encoding focus, judging from the perception results.

The results of the perception tests demonstrate the importance of post-focus compression for effective encoding of focus. This is seen in the fact that without such compression in Taiwanese and Taiwan Mandarin, focus recognition rate is

substantially lower than in Beijing Mandarin, and in the fact that even for Beijing Mandarin, when such compression is not possible in final focus, the recognition rate is similar to Taiwanese and Taiwan Mandarin.

Table 3. *Confusion matrix of focus perception (%) by 4 groups of listeners listening to utterances by their own language group. Bold face indicates correct focus identification.*

Listener group	heard as	None	Initial	Medial	Final
	original				
Taiwanese	None	50.7	22.0	16.0	10.7
	Initial	29.3	59.3	5.3	6.7
	Medial	22.0	19.3	45.3	14.0
	Final	30.7	6.7	8.7	54.7
Taiwan Mandarin	None	66.0	13.3	18.0	2.7
	Initial	32.0	63.3	4.0	0.7
	Medial	18.7	4.7	73.3	3.3
	Final	21.3	2.7	8.0	68.0
Bilingual Taiwanese	None	71.3	14.7	10.7	3.3
	Initial	27.3	61.3	10.0	1.3
	Medial	18.0	2.7	73.3	6.0
	Final	23.3	0.7	8.0	68.0
Bilingual Mandarin	None	72.0	11.3	6.7	10.0
	Initial	35.3	52.0	9.3	3.3
	Medial	25.3	6.0	63.3	5.3
	Final	32.0	4.7	10.7	52.7
Beijing Mandarin	None	78.8	6.7	7.3	7.3
	Initial	7.9	90.9	1.2	0
	Medial	6.7	0.6	92.7	0
	Final	27.9	0	5.5	66.7

One of the most unexpected outcomes of the present study is the finding that Taiwan Mandarin, which is closely related to Beijing Mandarin, realizes focus in a manner very similar to that of Taiwanese. Taiwan Mandarin was originally included in the study as a control for highlighting the focal differences between Mandarin and Taiwanese. To our surprise, its focus realization turned out to be much more similar to Taiwanese than to Beijing Mandarin. This assimilation may have been the result of an intimate contact between the two languages in over 60 years [12]. An important form of this intimate contact is through bilingualism, which has been highly common in Taiwan in the past decades [12].

That post-focus compression as a focus encoding strategy can be either present or absent in languages/dialects as similar as Beijing and Taiwan Mandarin suggests that its adoption is independent of the tonal typology of the language. The “loss” of such encoding strategy by Taiwan Mandarin through its contact with Taiwanese suggests that its presence/absence is more closely related to historical changes through language contact. Such a hypothesis of course can be verified only by examining many more languages of the world.

4. Conclusions

The present findings show that even very closely related languages can have rather different ways of realizing focus. Such difference is independent of whether the language is tonal, as Taiwanese, Taiwan Mandarin and Beijing Mandarin are all tonal, or whether there are morphosyntactic means to indicate focus, as they exist in both Mandarin and Taiwanese. The case of Taiwan Mandarin is especially revealing, as it is phonetically very similar to Beijing Mandarin, and yet its focus realization is more similar to Taiwanese, with which it

has been in close contact for several generations.

The present data have once more demonstrated the perceptual importance of post-focus compression — reducing the pitch range and intensity of post focus words. The presence of such compression in initial and medial focus in Mandarin lead to over 90% focus recognition, whereas the lack of it in final focus in Beijing Mandarin and in all types of focus in Taiwanese and Taiwan Mandarin lead to less than 75% of focus recognition.

Our findings thus call for large scale typological investigations of the world’s languages to test the hypothesis that post-focus compression spreads through language contact rather than arising automatically due to the tonal characteristics of each language. Our data also suggest the importance of using systematic experimental control in such investigations, including eliciting focus with context, using identical target sentences, taking measurements from both on-focus and off-focus syllables, and the inclusion of a no focus condition as the base line for comparison.

5. Acknowledgements

The research is supported in part by NIH grant No. 1R01DC006243 to the third author.

6. References

- [1] Boersma, P., “Praat, a system for doing phonetics by computer,” *Glott International* 5:9/10:341-345, 2001.
- [2] Botinis, A., Fourakis, M., and Gawronska, B., “Focus identification in English, Greek and Swedish,” in *Proceedings of The 14th International Congress of Phonetic Sciences*, San Francisco, 1557-1560, 1999.
- [3] Chen, M. Y., *Tone Sandhi: Patterns across Chinese Dialects*, Cambridge, UK: Cambridge University Press, 2000.
- [4] Cheng, Robert L. “A Comparison of Taiwanese, Taiwan Mandarin, and Peking Mandarin”. *Language* 61:352-377, 1985.
- [5] Fon, J., Chiang, W.-Y., & Cheung, C. Production and perception of two dipping tones (T2 and T3) in Taiwan Mandarin. *Journal of Chinese Linguistics*, 32: 249–280, 2004.
- [6] Ishihara, S., “Major phrase, focus intonation, and multiple spell-out (MaP, FI, MSO),” *The Linguistic Review* 24:137-167, 2007.
- [7] Kügler, F., and Skopeteas, S., “On the universality of prosodic reflexes of contrast: The case of Yucatec Maya,” in *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany, 2007.
- [8] Liu, F., and Xu, Y., “Parallel encoding of focus and interrogative meaning in Mandarin intonation,” *Phonetica* 62:70-87, 2005.
- [9] Pan, H., “Focus and Taiwanese unchecked tones,” *Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation*, in C. Lee, M. Gordon and D. Büring, [eds], 195-213, Springer, 2007.
- [10] Pierrehumbert, J., and Beckman, M., *Japanese Tone Structure*, Cambridge, MA: The MIT Press, 1988.
- [11] Rump, H. H., and Collier, R., “Focus conditions and the prominence of pitch-accented syllables,” *Language and Speech* 39:1-17, 1996.
- [12] 曹逢甫(Tsao, Feng-fu), *臺式日語與臺灣國語：百年來在臺灣發生的兩個語言接觸實例*。漢學研究: 273-97, 2000.
- [13] Xu, Y., “Effects of tone and focus on the formation and alignment of F₀ contours,” *Journal of Phonetics* 27:55-105, 1999.
- [14] Xu, Y. “_TimeNormalizeF0.praat. Online: <<http://www.phon.ucl.ac.uk/home/yi/tools.html>>,” 2005-2009.
- [15] Xu, Y., and Xu, C. X., “Phonetic realization of focus in English declarative intonation,” *Journal of Phonetics* 33:159-197, 2005.
- [16] Xu, Y., Xu, C. X., and Sun, X., “On the Temporal Domain of Focus,” in *Proceedings of International Conference on Speech Prosody 2004*, 81-84, 2004.
- [17] Zheng, Chin-chin. “A Contrastive Analysis on the Pronunciation between Mainland China and Taiwan: A Case Study of News Reporters.” Taipei: NSC-88-2411-H-003-017, 1999.